Philosophia Scientiæ
Travaux d'histoire et de philosophie des sciences

9-1 | 2005
Varia

Counterfactuals: The Epistemic Analysis

John-Michael Kuczynski

Electronic version
URL: http://journals.openedition.org/philosophiascientiae/618
DOI: 10.4000/philosophiascientiae.618
ISSN: 1775-4283

Publisher
Éditions Kimé

Printed version
Date of publication: 1 May 2005
Number of pages: 83-126
ISBN: 2-84174-368-3
ISSN: 1281-2463

Electronic reference
John-Michael Kuczynski, « Counterfactuals: The Epistemic Analysis », *Philosophia Scientiæ* [Online], 9-1 | 2005, Online since 24 June 2011, connection on 30 April 2019. URL : http://journals.openedition.org/philosophiascientiae/618 ; DOI : 10.4000/philosophiascientiae.618

Counterfactuals : The Epistemic Analysis[1]

John-Michael Kuczynski
University of California, Santa Barbara

Résumé : En temps normal, les contrefactuels sont conçus comme produisant des énoncés portant sur des états de choses, mais des états de choses se trouvant dans des mondes simplement possibles ou alternes. Analysés ainsi, il s'avère que presque tous les contrefactuels sont incohérents. Tout contrefactuel analysé de la sorte exige qu'il y ait un monde métaphysiquement (et pas épistémiquement seulement) possible w où les lois sont les mêmes qu'ici, et où la quasi-totalité des faits sont les mêmes qu'ici. (Les différences factuelles sont liées à l'antécédent et au conséquent du contrefactuel). Mais comme je vais le montrer, cette exigence implique de manière typique la postulation de mondes dont on peut démontrer la nécessaire non-existence au moyen de déductions assez élémentaires. De plus, l'analyse des contrefactuels en termes de mondes possibles est coupable de circularité cachée. Analysés ainsi, les contrefactuels ne peuvent effectivement être compris qu'en termes de lois de la nature (les lois qui s'appliquent ici sont présupposées dans le monde hypothétique — excepté dans le cas atypique où le contrefactuel est également un contre-nomique). Mais il y a un argument répandu voulant que le concept de loi ne puisse pas être défini lui-même, sinon dans les termes de la notion de contrefactuel (une loi est, entre autres, quelque chose qui vient à l'appui des contrefactuels). Je donnerai une analyse purement épistémique des contrefactuels, déclarant que ce sont des propositions pseudo-probabilistes. Je déclarerai également que le type de probabilité en question peut être défini intégralement en termes de ce qui a été le cas (non pas de ce qui aurait été le cas, ni même de ce qui doit être le cas en un sens nomique.) Mon analyse n'est donc coupable d'aucune sorte de circularité.

[1]This paper benefited immeasurably from the incisive points made by an anonymous reviewer at *Philosophia Scientiae*. Nearly all of the points attributed to the imaginary interlocutor are points made by the reviewer. I would also like to thank Ernan McMullen for reading an earlier draft.

Abstract: Ordinarily counterfactuals are seen as making statements about states of affairs, albeit ones that hold in merely possible or alternative worlds. Thus analyzed, nearly all counterfactuals turn out to be incoherent. Any counterfactual, thus analyzed, requires that there be a metaphysically (not just epistemically) possible world w where the laws are the same as here, and where *almost* all of the facts are the same as here. (The factual differences relate to the antecedent and consequent of the counter-factual.) But, as I show, this requirement typically involves the positing of worlds whose necessary non-existence can be shown by fairly elementary deductions. Further, the possible-worlds analysis of counterfactuals is guilty of covert circularity. For, thus analyzed, counterfactuals can only be understood in terms of laws of nature (the laws that apply here are assumed in the hypothetical world - except in the atypical case where the counterfactual is also a counter-nomic). But there is widespread agreement that the concept of a law cannot itself be defined except in terms of the notion of a counterfactual (a law is *inter alia* something that supports counterfactuals). I give a purely epistemic analysis of counterfactuals, arguing that they are crypto-probability propositions. I also argue that the relevant kind of probability can be defined wholly in terms of what *has* happened (not what *would* happen and not even what must happen in a nomic sense). So my analysis isn't guilty of any kind of circularity.

I. Statements like the following are examples of *counterfactuals* :

(1) If Kennedy had not been assassinated, the U.S. would have pulled out of Viet Nam before 1965.

(2) If Jones had not fired the gun, Smith would still be alive today.

A counterfactual says what *would* have been the case *if* some condition C obtained, when in fact *C did not* obtain.

The consensus is that no acceptable analysis of counterfactuals has been proposed. I agree with the consensus. The reason there is no satisfactory analysis is that the existing analyses regard counterfactuals as saying something about causal relations among *states of affairs* — albeit ones that take place in non-actual or actual but alternative worlds. The right approach, I argue, is to see counterfactuals as expressing relations of probabilification holding among *statements*. Counterfactuals don't really say anything about *events* or anything concrete — or even anything that *would* be concrete if it existed. Counterfactuals merely register confirmation relations among propositions. That is what I will try to show in this paper.

I must make one thing clear. I am taking it for granted that counterfactuals are, or can be, true. I am not arguing that they are necessarily incoherent or otherwise false. So *given* that a factualist analysis makes them incoherent, a non-factualist analysis — an epistemic one — is needed.

Of course, it has been argued (by Quine, for example in [Quine 1982, 23]) that counterfactuals *are* meaningless or, at least, systematically false. But given how endemic they are to our thought, it would be better to find a way to make them non-meaningless than to try to do without them[2]. (I am open to the idea that they are all meaningless or false, but I will try to show how we can avoid coming to that undesirable conclusion.)

II. David Lewis has given the most perspicuous account of the conceit that underlies ordinary thinking about counterfactuals, and that underlies virtually all philosophical thinking about it :

> "If kangaroos had no tails, they would topple over" seems to me to mean something like this : in any possible state of

[2]Clarence Lewis [Lewis 1929] argued compellingly that *every* statement about the physical world — even those that seem to have no counter-factual component — are in fact implicitly counterfactual. When you say that something is steel, you are saying something about how it *would* behave *if* it were subjected to certain conditions. I am inclined to agree with Lewis. There is no way to purge our discourse of counter-factual elements. So we have to find a way to make counterfactuals work.

> affairs in which kangaroos have no tails, and which resembles our actual state of affairs as much as having no tails permits it to, the kangaroos topple over (See [Lewis 1973, 1].

Notice that Lewis construes counterfactuals as making statements about *states of affairs*, albeit ones that hold in worlds other than this one. What makes (2) true is (inter alia) that, in some other world, Jones doesn't fire the gun and Smith still remains alive.

If Lewis is right, what (2) does *not* do is make a statement about logical relations among *propositions*. (2) is about facts following facts, bodies hitting (or not hitting) other bodies. If Lewis is right, (2) isn't about a shadow world of logical interrelations holding among platonic entities.

Let us describe as *factualist* any analysis that is, in this respect, like Lewis'. The factualist says that counterfactuals are about states of affairs, facts, and objects — the things we encounter in this world — except that they are in alternative worlds.

The factualist view is not tenable. That is what I will show. Counterfactuals do not say anything about relations among events, not even events that are in worlds other than this one. Counterfactuals say something *only* about the degree of credibility that certain propositions confer on other propositions.

III. Suppose that, in actual fact, Smith shoots Jones and thereby kills him. Given this, Let us consider the proposition :

(2) If Jones had not fired the gun, Smith would still be alive today.

This is a perfectly good counterfactual : the kind that might well be true. Let us apply the factualist analysis to it. According to the factualist, the meaning of (2) is :

(3) Consider a world w with the following two properties. First, in w, Jones doesn't pull the trigger. Second, with that qualification, w is as much like our world as it could be. In w, Smith is still alive.

One point must be made explicit. The *laws* obtaining in w must be the same as the laws that govern *our* universe. For if the physical laws of w are *different* from those in our world, then the counterfactual is immediately voided. If those laws are different, then Jones' pulling the trigger might not lead to the bullet leaving the gun, or Smith's flesh might repulse the bullet, or there might not even be a gun — perhaps those other laws wouldn't permit atoms to bind to one another in the right way. So clearly w must be governed by the same laws as our world.

Given this, let $L_1 \ldots L_n$ be the laws that govern our world. As we just saw, $L_1 \ldots L_n$ must also govern w. Now in our world, $C_1 \ldots C_k$ led to Jones' pulling the trigger. Of course, $C_1 \ldots C_k$ led to that event because laws $L_1 \ldots L_n$ govern our world. (A different set of laws needn't lead from those initial conditions to that particular conclusion.) Trivially, laws are either deterministic or they're not. A "deterministic" law is given by a proposition of the form : Given such and such conditions, thus and such happens *100%* of the time. An "indeterministic" law is given by a proposition of the form : given such and such conditions, thus and such happens *a certain percentage of the time* (not 100%). Let us first investigate the consequences of supposing that $L_1 \ldots L_n$ are deterministic. Then we will consider the consequences of supposing that at least *some* of the laws in $L_1 \ldots L_n$ are indeterministic. (This exhausts all the possibilities : either *all* of $L_1 \ldots L_n$ are deterministic, or some of them are *not* deterministic, i.e. are indeterministic.)

IV. *First case : suppose $L_1 \ldots L_n$ are deterministic.* In that case, in w, $C_1 \ldots C_k$ lead to Jones pulling the trigger. Now remember that we are trying to keep things in w *as much like things in our world* as is possible given that, in w, Jones did not pull the trigger. We've just seen that we cannot keep $C_1 \ldots C_k$ the same ; at least one of them must go. So, for some i, C_i must *not* be the case in w. But C_i did not just come out of nowhere ; some *previous* set of conditions $C_1^* \ldots C_m^*$ are what led to C_i. So given that C_i is *not* the case in w, and given that the laws leading from $C_1^* \ldots C_m^*$ to C_i are deterministic, we must suppose that, for some j, C_j^* is not the case. Otherwise we are stuck with C_i, and if we are stuck with C_i, we are stuck with Jones' pulling the trigger. Of course, C_j^* didn't just come out of nowhere ; it arose from some prior conditions $C_1^{**} \ldots C_o^{**}$. So we must assume, for now familiar reasons, that for some k, C_k^{**} is not the case. Clearly this regress goes on and on.

In fact, this regress goes back into the furthest recesses of the past. So it is clear that, if Jones is to *not* pull the trigger in w, then w's past must be *very* different from the past of our world. But if w's past is *that* different, then there is no guarantee at all that w's present would even *remotely* resemble our present. What we *want* is for w's present to be just like ours *except* that Jones doesn't pull the trigger. But what we get is a present which almost certainly doesn't even have a Jones or a Smith and which, even if it does comprise Smith and Jones, is so different from our world that nothing can be said with confidence about it.

In a deterministic world, any event is deeply embedded in other events. Consequently, two deterministic worlds cannot differ with respect to a single event unless they differ with respect to countless other events. Given that Jones didn't *actually* shoot Smith, a world that is different enough from ours to have Jones shooting Smith will, paradoxically, be *too* different to include that event. And even if (what is virtually impossible) Jones *does* shoot Smith in w, the conditions that surround that shooting will probably be so different that, for all we know, Smith *does* survive. If the conditions in w are different enough to allow Jones to shoot Smith, then they may be different enough to allow Jones to be wearing a bullet-proof vest.

So if the factualist approach is to work, then our world must be *indeterministic*. Let us now consider this option.

V. *Second case : suppose some laws are indeterministic*. Once again, let $L_1 \ldots L_n$ be the laws governing our world. And suppose that, for some i, L_i is indeterministic.

First of all, counterfactuals are only as good as our laws are deterministic. Suppose that being shot in the chest is fatal 100% of the time. In other words, suppose the law : *x is shot in the chest*$\rightarrow$ *x dies* is deterministic. In that case, I can say with complete certainty "if Jones *had* been shot in the chest, he *would* have died".

But if that law is only probabilistic, then the counterfactual is not true. If being shot in the chest leads to death 39% of the time, or even 72% of the time, then the counterfactual "if x had been shot in the chest, x would have died" is definitely not *true*. It seems to lack truth-value. What *can* truly be said is "if x had been shot in the chest, then there would have been a 39% (or 72%) of x's dying." But that is an entirely different statement from the original counterfactual.

Of course, given only that *some* laws are indeterministic, it doesn't follow that *all* counterfactuals are thereby vitiated : only those counterfactuals are vitiated which depend on the probabilistic laws in question.

Given this, it might seem that if *some*, but not all, laws are indeterministic, then the factualist position might work.

But it is easy to show that the existence of *some* indeterministic laws doesn't help the factualist at all. It makes things worse for him.

To help put the reasons for this in context, let us momentarily forget that we are talking about *in*deterministic laws and let us once again talk about deterministic laws. Let us once again suppose that $L_1 \ldots L_n$ (the laws that govern our universe) are *deterministic*. It is, at the very least,

an epistemic possibility that this supposition should turn out to be true. Taking 'might' in a purely epistemic sense, future science might reveal that all our laws are deterministic.

If this should turn out to be the case, it would obviously *not* mean that all our counterfactuals statements would turn out to be false. Consider the following counterfactual :

(4) If Brown had struck the window with the hammer, it would have broken.

Let us suppose that the window is very thin, that Brown is strong, and so on. In other words, let us suppose that (4) is a true counterfactual. If future physicists discover that quantum mechanics is deterministic, would it mean that (4) turns out to be false ? Suppose that in the year 2135 it is finally established that all physical laws are deterministic. Will people then say : "I guess we were wrong to say that *if* President Clinton had resigned, Gore would have become President" ? Of course not. Those counterfactuals would still be true.

An *analysis* of counterfactuals cannot make the truth of such statements depend on things on which their truth is not *in fact* dependent. It is a datum that (4) is true. It is not a datum (though it may be true) that

(5) there are indeterministic laws.

So the truth of (4) does not depend on that of (5). But the factualist's analysis makes the truth of (4) depend on that of (5). So the factualist's analysis is wrong.

The specificity of the factualist's demands

There is more to say. If indeterminism is to make the factualist analysis work, it is not enough that world be indeterministic in *any* way. On the contrary, it must be indeterministic in extremely specific ways. The factualist needs the indeterminism to be operative in specific contexts, and in specific ways *within* those contexts. At the same time, the factualist needs other contexts to be resolutely deterministic.

Consider the laws mediating between window-strikings and window-shatterings. If *those* laws are indeterministic, then (4) comes out false. The indeterministic character of the laws involved might allow the window to stay intact. So if (4) is to come out true, then those laws must be resolutely deterministic.

At the same time, the factualist needs *in*determinism to prevail in other contexts. This is a consequence of what we said earlier. The factualist says that the occurrences in some world w are what make (4) true. w must be governed by the same laws as our world. If our world is

too deterministic, then a single difference between our world and w will mean a *massive* difference between them. And, as we saw earlier, if there is such a massive difference, then the occurrences in w will not make (4) be true.

So the factualist-indeterminist (i.e. somebody who is a factualist but predicates the truth of counterfactuals on the existence of indeterministic laws) ends up being committed to extremely specific beliefs as to where indeterministic laws are operative. He needs the world to be deterministic in very specific ways, and he also needs the world to be indeterministic in very specific ways. It is *at best* an open empirical matter whether the world has just the right balance of determinism and indeterminism. (At worst, what the factualist wants is totally incoherent. I will discuss this in a moment.)

So if the factualist is right, then counter-factuals like (4) would turn out to be false if scientists discovered, as they well may, that the world doesn't have the right balance of determinism and indeterminism. But surely the truth of (4) is *not* contingent on the outcome of such inquiries. So the factualist has mis-analyzed (4).

But even this is too charitable to the factualist. On the factualist account, *different* mixtures of determinism and indeterminism are needed to validate different counter-factuals. The mixture of determinism and indeterminism needed to validate the factualist analysis of "if Jones had struck the window, the window would have shattered" is going to be different from the mixture needed to validate "if Jones hadn't taken his pills this morning, he would now very ill". Given that Jones *did* take his pills, certain laws must be indeterministic in a world w where Jones did not take his pills, but is otherwise like our world. Given that Jones *did not* strike the window, certain laws must be suspended in a world w^* where Jones *did* strike the window, but is otherwise like our world. But the set of laws that are indeterministic in w will inevitably not quite coincide with the set of laws that are indeterministic in w^*. Both w and w^* must be governed by the same laws as our world; otherwise occurrences in those worlds don't validate counter-factuals affirmed here. It follows that in *our* world, certain laws must be simultaneously deterministic and indeterministic. But this is absurd.

The factualist needs the world to be simultaneously characterized by *different* mixtures of determinism and indeterminism. But this is an incoherent requirement.

An objection

I will end this section by addressing a possible objection (objections to my analysis will henceforth be put in the words of an imaginary interlocutor) :

> The argument you just gave seems to involve the premise that the macroscopic world is "vanishingly close" to being deterministic. Admittedly, there is a fair amount of empirical support for this view, and it is a standard one. But it might turn out to be wrong. In that case, it seems, your argument would fail. So your argument presupposes the truth of a dubious empirical doctrine[3].

For the sake of argument, let's suppose I *am* wrong. Let's suppose that the macroscopic world is non-negligibly indeterministic. In that case, the factualist's position is automatically nullified. Suppose that, in our world, the exact same initial conditions, down to the last proton, lead to the car starting only 54% of the time. (Suppose that my turning the key in the ignition is, or eventuates in, one these initial conditions.) In light of this, consider the counter-factual :

(6) *If I'd turned the key, the car would have started*

What does the factualist say about (6) ? He says :

> Let w be a world just like ours with this one difference : in w, you *did* turn the key. With that qualification w is as much like our world as is possible. In w, the car *does* start.

But that isn't good reasoning at all. w must be governed by the same laws as here. So the same indeterminism that infects our macro-world must infect w's macro-world. Given this, it is obviously quite false to say that, in w, the car definitely *does* start. The most one can say is : there is a 54% chance that it starts in w. Really, the proposition *in* w, *under the circumstances described, the car starts* seems to lack truth-value. In any case, it definitely isn't true. So in so far as the macro-world here is indeterministic, then the factualist position is quite definitely false.

Summary

The factualist needs there to be worlds that are *almost* exactly like our worlds. For every proposition P that is not true of our world, the factualist needs there to be a world w with the following property : in w, P is true ; but apart from that, there are no significant differences between our world and w.

[3]This point was made to me by an anonymous reviewer at *Philosophia Scientiae*.

But either this demand cannot be met *or*, if it can be met, the conditions that enable it to be met make w incapable of validating the counter-factuals it is supposed to validate.

On the factualist account, when we run a counter-factual, we are considering a world w where the *laws* are the same as they are in our world.

Either our world is deterministic or it is not. Let us consider each option. First let us suppose our world is deterministic.

Two deterministic worlds cannot differ from each other with respect to a single state of affairs unless they differ with respect to a series of states of affairs extending back into the furthest recesses of time. Of course, if the pasts of two worlds differ that much, then their presents will differ in unfathomable ways.

So if Jones shot Smith in w, and Jones didn't shoot Smith in this world, then this world and w will differ in unfathomable ways. We thus have absolutely no idea what the consequences of Jones' shooting Smith in w might be. In fact, the chances of there even being a Jones or a Smith in w are infinitesimal. So the occurrences in w are useless when it comes to validating counter-factuals.

So if our world is deterministic, then the factualist approach is a failure. So let us suppose that our world is indeterministic.

This supposition destroys the factualist's position. Counter-factuals are sustainable only in so far as the world is deterministic. So to the extent that counter-factuals *can* be truly affirmed, the world is deterministic. But then we run into the problem just described.

So the factualist analysis fails if the world is deterministic, and it fails if the world is not deterministic. So it fails.

VI. But we have not yet completely refuted the factualist's position. For there is one move that the factualist might still make[4] :

> Suppose that in actuality, Jones does not shoot Smith. And suppose that Jones is psychologically normal and thus *would* feel remorse if he did something as heinous as shooting somebody.
>
> Given this, consider the counter-factual :
>
> (7) If Jones had shot Smith, Jones would now be feeling profound remorse.

[4]Lewis in [Lewis 1973] makes this move.

> Let w be some world that is just like ours except for this one thing. In w, at time t, Jones shoots Smith. In the actual world, at time t, Jones does not shoot Smith.
>
> I grant that the factualist runs into the problem you described *if* he insists on tracing every difference between w and our world to some pre-existing difference between them. For then, obviously, he will have to suppose that w's entire history deviates from ours.
>
> But there is no reason why the factualist must do that. The factualist can say this. Suppose that *up until* t w and the actual world are exactly the same. But at t, the two worlds diverge. In w, Jones shoots Smith. In our world, Jones does not Smith.
>
> Of course, if w is to diverge from the actual world in this way, then some miracle must occur in w. For given the laws and initial conditions that obtain in our world, Jones *must* shoot Smith at t. Thus, Jones *must* shoot Smith at t in w as well. But this 'must' can be neutralized by a miracle. (There are metaphysically possible worlds where miracles happen.) And this is just what happens in w at t. Thus w differs from the actual world in *just* the one respect the factualist requires. Consequently w can validate the counter-factual in question in just the right way.

The literature on counter-factuals is replete with moves like the one just described. But that move fails on multiple levels. It involves profound but easily demonstrated misunderstandings concerning both the nature of persistence throughout time and also concerning the nature of natural law.

According to the factualist, our world is just like w *until* t : same laws, same conditions. But, in w, some miracle happens at t. That is why, in w, Jones manages to shoot Smith.

The first thing to point out is that it is not just *one* miracle that happens. It is billions upon billions. Shooting somebody is not a simple event. Various psychological, metabolic, cellular, molecular, atomic, and sub-atomic events are involved. Your pulling a trigger involves innumerable bio-chemical and, therefore, atomic events. The miracle in question involves *each* of those atoms and protons and quarks occupying a place that, nomically, it is not supposed to occupy. Countless particles must suddenly be re-directed or, failing that, be annihilated. This would involve suspending many, perhaps all, of the laws involved the behavior of

those particles. So *each* atom-relocation involves a miracle. In fact, each such relocation involves *several* miracles. For more than one physical law will have to be suspended in each case.

Let Alpha be any one of the particles composing Jones, and let t be the time of the miracle. The miracle in question involves severing many, probably all, of the causal ties that link Alpha–before–t to Alpha–after–t.

Now we can begin to see the incoherence in what the factualist is demanding. Objects just *are* causal sequences. They are series of states of affairs that are linked together by physical laws. If x and y have no causal connections to each other — if they are not installments in the same causal process — then x and y are not the same object.

The factualist wants us to suspend the laws that link Alpha–before–t to Alpha-after-t. So he wants us to sever the causal connections between the former and the latter. If he severs *all* such connections, then Alpha-before-t and Alpha–after–t are simply different entities. The factualist has simply *annihilated* a certain particle, not redirected it.

The laws that *govern* the behavior of an entity are *required* for its continued existence. Causal ties do not *bind* an object; they *compose* it. If you sever them, the object isn't liberated. It ceases to exist.

The factualist needs to sever the ties that bind Jones-before-t to Jones-after-t. But if he does this, then he has simply annihilated Jones, not redirected him. So if we grant the factualist his miracle, then there is no Jones after t. *A fortiori* there is no Jones who feels remorse.

Of course, the factualist could argue that the miracle in question doesn't sever *all* of the ties that are involved in Jones' continued existence. Some of those ties — the ones needed to sustain Jones' existence — are left unsevered. But other ties — the ones that, in our world, lead Jones to do something other than shoot Smith — are severed.

But it is exceedingly unlikely that such a neat separation can be effected between those two sets of causal ties. *If* we are going to redirect all of those atoms, molecules, cells, and so on, then we must put all manner of physical laws out of commission. The laws of electrodynamics, mechanics, thermodynamics, and so on, must all be suspended. But if all of these laws are suspended, what is left over that can keep Jones in existence? Jones is a causal process. In all likelihood, every law of physics will have to be suspended to create the miracle that the factualist wants. Remember the factualist is asking for *billions* of miracles; he is asking *billions* of particles to *not* do what, as a matter of physical law, they must do. Leaving aside highly derivative laws like *pinewood floats on water*, it is exceedingly unlikely that any law of physics will escape the

factualist's wrath.

The needed cut in causal ties is so deep that what is left over, if anything, will almost certainly not be enough to sustain Jones' existence. As for the chances of Jones' continuing to exist in good enough physical and psychological *health* to feel remorse — those chances are essentially zero.

The non-containability of miracles

There is another reason to hold that miracles do *not* help the factualist, indeed, that they have no place in discussion of counter-factuals.

A miracle is a suspension of a natural law. But it isn't possible to contain or isolate suspensions of causal laws. So if we help ourselves to a miracle, then we are committed to all manner of consequences that are not even counter-factually true.

Suppose we don't want some event E_1 to occur in our counterfactual scenario; and we keep E_1 out by suspending some set of laws $L_1 \ldots L_n$ What ends up happening is that, because of that suspension, various *other* events $E_2 \ldots E_k$ that we *don't* want to happen end up happening. So in suspending $L_1 \ldots L_n$, we end up committing ourselves to counterfactuals that are not true or that, in any case, are very different from the one we want to assert.

Once again, consider (7). The laws and initial conditions that prevailed in our world required Jones *not* to shoot Smith. Those same laws and conditions obtain in w. So if Jones is to shoot Smith in w, there must be a miracle : the laws binding Jones to *not* shooting Smith must be suspended. But if those laws are put out of operation, even for a moment, the result is completely and utterly unpredictable. It is as likely that Jones turns to gelatin as it is that he feels remorse.

This becomes apparent when we start to think more rigorously about the nature of causation. We say that *the* cause of the crash was Brown's inattention to the road. But the truth is that innumerable conditions had to obtain for Brown to crash. The road and the tires had to be made of a certain material; the tread on the tires had to be worn down by a certain amount; the road had to be wet; the car had to be moving with a certain speed; the mechanisms connecting the steering wheel to the rotations of the tires had to have, or lack, a certain degree of integrity; and on and on[5]. Given virtually any event, the conditions that were causally necessary for that event to occur are innumerable[6].

[5]This example was borrowed from Carnap, who makes this very point.

[6]Indeed, an old debate in the philosophy of science — one that has still not been settled (but is no longer focused on) — is the question whether causal relations hold

The other side of the coin is that, if some set of conditions $C_1 \ldots C_m$ is sufficient to bring about a given event, then it is *also* sufficient to bring about innumerable other events. Nothing that causes anything causes just *one* thing.

Given this, suppose that, in w, $L_1 \ldots L_n$ are suspended, so that $C_1 \ldots C_m$ don't lead to Jones' pulling the trigger. That suspension will not *just* lead to Jones' not pulling the trigger; it will have innumerable other consequences. It is overwhelmingly likely that what was necessary to get Jones to pull the trigger was *also* necessary for Jones' head not to explode, for the house he was in not to disintegrate, for the electrical activity in Jones' brain not to cease. Among the necessary conditions of Jones' pulling the trigger — in other words, among $C_1 \ldots C_m$ — are facts having to do with the integrity of the various alloys composing the gun; facts relating to the architecture of the house Jones was in; facts relating to the durability of the materials composing the floor Jones' was standing on; facts relating to the amount of pressure that Jones' vertebrae could withstand without pulverizing; and so on. If we suspend $L_1 \ldots L_n$ just enough to make sure that Jones does not pull the trigger, then we are suspending them quite enough to vaporize Jones' spinal cord, to melt Smith's brain, to turn the house that they are both occupying into a heap of smoldering rubble. So if we analyze (7) as making a statement about some world where there is a miracle, then we cannot hold (7) without *also* holding some absurd proposition like :

(8) If Jones hadn't pulled the trigger, then the house Smith and Jones are occupying would have turned to dust

For if we suspend the laws in w enough to get Jones to shoot Smith, we will be suspending them enough to make all manner of other, unwanted things happen.

The factualist needs to use miracles. But miracles turn counter-factuals into counter-*nomics*. So the factualist turns counter-factuals into statements about occurrences in worlds that are not, strictly speaking, governed by the same laws as ours. Given two worlds that differ in respect of their laws, even in the most minor way, the differences between them are incalculable. A counter-nomic scenario will differ *so* much from the actual world that it couldn't possibly validate any counter-factuals, at least not without simultaneously validating many false ones.

---

between the state of the entire universe at one time and the state of the entire universe at a later time. See the last chapter of Russell's *Principles of Mathematics*. If causal relations hold between states of the universe at different times — and it is still an open question whether they do — then the factualist position is immediately seen to be untenable.

The incoherence of the concept of a miracle

So far we have allowed that miracles *are* in some sense possible. Philosophers writing about counter-factuals make practically unrestricted use of miracles. So do ordinary people in their day to day thinking. It is generally taken for granted that miracles are *in some sense* possible.

But the concept of a miracle is one that has no utility at all for the factualist. The very concept of a miracle is incoherent *unless* a quite specific, and highly controversial, conception of natural law is the right one. But *if* that conception of law is the right one, then the correct analysis of counter-factuals is absolutely *not* the factualist one.

Ultimately, there are two ways of thinking about laws. On the one hand, laws can be thought of as mere regularities. This is how Hume thought about them. On the other hand, they can be thought of as things that *ground* regularities — as things that *guarantee* that, when such and such happens, thus and such will also happen.

Let us consider each of these conceptions, and see what the consequences are for the factualist.

Non-Humean miracles

First, suppose that Hume is wrong about laws; let us suppose that laws *ground* regularities, and are not *identical* with them. In that case, as a matter of conceptual necessity, laws are *not* the kinds of things that can be suspended.

Imagine the following. Some state of affairs $\underline{a}$ causes some other state of affairs $\underline{b}$. The relevant covering law is : If there is an A at t, then there is a B at t^*.

Given that $\underline{a}$ really is the cause of $\underline{b}$, it follows trivially that $\underline{a}$ *by itself* guarantees $\underline{b}$'s existence. It isn't as though something must be *added* to $\underline{a}$ to make $\underline{b}$ happen. If some such addition had to be made, then contrary to our supposition, $\underline{a}$ would *not* be the cause of $\underline{b}$. It would, at most, be a partial cause.

This means that *merely* by virtue of the fact that we have $\underline{a}$, we have enough for $\underline{b}$. In other words, given *only* that we have an instance of type A, we have enough for an instance of type B.

This, in turn, means that if the antecedent of some law is fulfilled, then *ipso facto* that law is instantiated and is operative. If we have an A, then *ipso facto* the law that generates a B is instantiated.

The only way you can have an A *without* having a B is if A is not *by itself* sufficient for B. But if A were thus deficient, it wouldn't cause B at all. It would be, at most, a partial cause.

Law are not things *added* to states of affairs. They are internal to the states of affairs they govern. You cannot remove the laws without removing the states of affairs themselves. Such a removal cannot even be thought, at least not coherently. Of course, such removals *are* thought. But those thoughts involve the absurd idea that something must be *added* to a causally sufficient condition to *make* it be a causally sufficient condition.

So if we think of laws as things that tie events together, and not just as regularities, then laws can no more be suspended or removed than triangles can have four sides.

Humean miracles

So if the factualist is to have his miracles, then he must *not* think of laws as things that tie events together. So he must be a Humean about laws ; he must see laws as mere regularities.

But if laws are Humean, then the possible-worlds approach to counter-factuals is completely bankrupt.

The Humean says that laws are *just* regularities. For A's to cause B's is *just* for B's to always come right after A's. There is no inner connection ; nothing *grounds* the regularity. The regularity is all that is there. Nothing *makes* the regularity hold ; it just holds.

Given this, let w and w^* be two worlds that are exactly the same, in every conceivable respect, up until time t. The Humean says : there is absolutely no reason why w and w^* should not diverge after t. The Humean actually allows for *any* divergence at all. The divergence can be total.

Why does the Humean allow for such a divergence ? The Humean doesn't countenance the idea of there being *forces* of any kind. Nothing *makes* anything else happen. Forces, if they exist at all, are just regularities. There are no nomic necessities. There is nothing to a world *underlying* the states of affairs in it. So no matter how much two worlds are alike before t, there is no limit to how much they can differ — not just logically but *nomically* — after t.

This means that, if you are a Humean, alternative worlds are useless for validating counter-factuals. Suppose, once again, that in our world, Bob did not shoot Smith at t. And suppose that w is some other world where Bob *does* shoot Smith at t, but is otherwise just like our world. (To make things as easy as possible for the factualist, let us waive all the problems we've been discussing relating to this kind of supposition.) A Humean will say :

> I have absolutely no idea what happens in w after t. It is perfectly possible that Smith feels no remorse ; it is perfectly possible that he turns into a blue-berry. w doesn't favor the one outcome more than the other.

Laws are just regularities. Events just happen. There are no underlying *dispositions*. In our world, it happens that people often feel remorse after committing crimes. But this fact is not *grounded* in some underlying disposition. Our world is no *more* disposed to favor Smith's feeling remorse after shooting Jones than it is disposed to favor Smith's turning into a blueberry. It just happens that Smith *does* feel remorse and *does not* turn into a blueberry.

By the same logic, w is no more disposed to favor Smith's feeling remorse after shooting Jones than his turning into a blueberry. It could well be that, in w, Smith *does* turn into a blueberry. Nothing about w favors one outcome more than any other. To think otherwise is to have a wrong conception of natural law : it is to think of laws as things that *underlie* regularities. In fact, laws *are* the regularities.

For a Humean, there will be as many worlds where Brown feels no remorse after shooting Smith — where, in fact, he turns into a blueberry or feels elated or succumbs to amnesia — as there are worlds where he does feel remorse. So, for a Humean, what goes on in other worlds is as likely to falsify a counter-factual as it is to validate it. Actually, if Hume is right, then — depending on one's views on the arithmetic of infinitely large numbers — what goes on in other worlds is *infinitely* more likely to falsify a given counter-factual claim than it is to validate it. For every world where Brown feels remorse after shooting Smith, there are infinitely many worlds where Smith doesn't feel any remorse.

To sum up, if laws are non-Humean, then they cannot be suspended. Miracles are not even *conceptually* possible. And the factualist's position fails. On the other hand, if laws *are* Humean, then the occurrences in other worlds are at least as likely to falsify our counter-factuals as they are to validate them. So miracles are useless for the factualist.

Some final remarks

There are two incidental points. We've seen that, if laws are non-Humean, then miracles are surds — they are in the same category as four-sided triangles. But it can be shown that, if laws *are* Humean, miracles are no better off.

If laws are Humean, then given such and such initial conditions, the world is no more disposed to do one thing than another. In a Hume-world, the world is no more disposed to produce a vase-shattering after

a vase-dropping than it is to produce a case of a vase turning into a pumpkin. So if, after being dropped from a tall building, the vase should turn into a pumpkin, that would be no more miraculous than if it were to shatter. Of course, if the vase shatters, there is no miracle. Thus, if the vase turns into a pumpkin — or simply fails to break — that is no miracle either. In a Hume world, *nothing* that happens, no matter how unusual, is a miracle.

So if laws are Humean, the concept of a miracle has no content. As we've seen, if laws are *non*-Humean, then that concept is incoherent. So the concept of a miracle seems to be quite broken.

The other incidental point is this. As we've just seen, for a Humean, two worlds can be exactly alike, in every conceivable respect, before time t, and yet diverge without limit after t. This has often been taken to show that Humean laws do not support counter-factuals (See [Mackie 1973].

But this is absolutely the wrong conclusion to draw. The right conclusion to draw is that Humean laws don't support counter-factuals *if* counter-factuals are seen as making statements about alternative worlds. If counter-factuals are understood in some other way, then there is a chance that Humean laws *do* support counter-factuals. We have seen that, indeed, counter-factuals *must* be understood in some other way : the alternative-worlds account is a failure. And we will see later that, in fact, Humean laws *do* support counter-factuals, contrary to what is generally held.

VII. Now it is time to give a positive analysis. As we've seen, counter-factuals *do not* make statements about states of affairs. Given this, there is, ultimately, only one thing to say. We must see counterfactuals as making purely *epistemic* statements — as making statements about relations between *propositions*. They must be statements to the effect that one proposition gives a very high degree of probability (perhaps 100%) to some other proposition.

Let us get clear on a couple of basic points about probability. First of all, probability is a relation between *propositions* ; it is not a property of states of affairs. Statements like the following are often true :

(9) The probability that Socrates had four kidneys is 5%.

But either Socrates did, or he did not, have four kidneys. It couldn't be 5% true that he had four kidneys. The same is true *mutatis mutandis* of

(10) The probability that Bill is on the North Pole right now is .002%

But obviously it cannot .002% true that Bill is on the North Pole. It is either a 100% true or 0% true.

Obviously (9) and (10) make statements about the credibility that certain propositions give to other propositions[7]. Let us focus on (9). Anyone affirming (9) is affirming it on the basis of certain pieces of (putative) knowledge that he has. He knows that most people don't have four kidneys, but that some people do. He knows, on the basis of a careful reading of the platonic dialogues, that Socrates had some of the symptoms had by people with that condition. But he also knows that Socrates had some other illness which would explain his having those symptoms.

Let $P_1 \ldots P_n$ be all these propositions. In affirming (9), our man is really saying that $P_1 \ldots P_n$ give a certain degree of credibility to the proposition that Socrates had four kidneys. Obviously the same considerations (mutatis mutandis) apply to (10).

So (9) is really elliptical for : $P_1 \ldots P_n$ *give such and such credibility to the proposition that Socrates had four kidneys*, where $P_1 \ldots P_n$ are the propositions our man believes true.

Of course, different people can token a sentence like "the probability that Socrates had four kidneys is 5%". And those people can have different data in mind. Under this circumstance, different statements are being made. Joe might token (9) on the basis of some piece of knowledge unknown to Jerry, who tokens the *same* sentence. We must hold that, in such a case, Joe and Jerry are really making different statements : Joe is saying

(11) *S gives such and such credibility to the proposition* Socrates is bald

where S is some conjunction of proposition.

On the other hand, Jerry is saying

(12) S^* *gives such and such credibility to the proposition* Socrates is bald

where S^* is some set of propositions other than S.

Consider the following dialogue :

> Smith : How probable is it that Brown watches over three hours of T.V. a day ?
> Jones : The probability is around 2%. For only 2% of human

[7]Keynes' work on probability is based on this point, which he very clearly states at the beginning of his treatise. See [Keynes 1921, §§ 2–3].

beings watch that much T.V.
Smith : But Brown is an American.
Jones : In that case, the probability is 75%.
Smith : But Jones is a prolific author.
Jones : In that case, the probability goes down to 4%. You see, prolific authors tend to watch very little T.V.
Smith : But Brown's writing concerns contemporary T.V. shows.
Jones : In that case, the probability goes up to 85%.

Each of Jones' statements is (or may be) correct. How is this possible ? For Jones seems to be contradicting himself. First he says that a certain proposition has only a 2% chance of true. But then he says it has a 75% chance of being true.

In fact, Jones is *not* contradicting himself. First he says : *given that Brown is a human, there is a 2% chance he watches more than three hours of T.V. a day.* Then Jones says : *given that Brown is an American, there is an 85% chance that he watches that much T.V.* These assertions are consistent.

It is clear, then, that probability is an epistemic, not a metaphysical, notion. Probability is a logical relation between propositions ; it is not a property of states of affairs.

When you utter (9), you are saying that, given the data at your disposal, it is *epistemically* possible to degree .05 that Socrates will turn out to have had four kidneys. In other words, $P_1 \ldots P_n$ give probability .05 to the proposition *Socrates had four kidneys*, where $P_1 \ldots P_n$ are the propositions you know to be true.

Obviously it is not *now* metaphysically possible for Socrates to have had four kidneys (in this world) unless he *actually* had four kidneys. *That* is an all or nothing affair — no percentages can be attached to a *truth*. So when you affirm (9), you couldn't possibly be talking about what is now metaphysically possible, but rather about what is now epistemically possible. And that means : what is not ruled out by the propositions that are currently known to you.

The meaning of

(13) if Kennedy hadn't died in '63, Johnson wouldn't have become President in '63

can only be

(14) The propositions $P_1 \ldots P_n$ give a high degree of credibility to the proposition *Johnson didn't become President in '63*,

where $P_1 \ldots P_n$ is some set of propositions that includes : *Kennedy didn't die in '63.*

It must be kept in mind that false propositions probabilify other propositions just like true ones. The proposition *the moon is made of cheese* is false, but that doesn't stop it from giving a very high degree of probability (essentially 100%) to the proposition *there is a piece of cheese that weighs more than 100 lbs.* The proposition *Kennedy didn't die in '63* is false, but it still gives probability to other propositions. It gives nearly 100% probability to the proposition *some living organism weighing over 10 lbs didn't die in '63.* And it gives some probability, though not nearly as much (at least not when taken by itself), to the proposition *Johnson didn't become President in '63.*

Notice that our purely propositional analysis of counterfactuals doesn't have the problem faced by the factualist analysis. The factualist analysis needs to keep the laws the same but vary the events : which, as we've seen, it is typically not *logically* possible to do without undermining the counterfactual. Given that it is *logically* impossible, it follows that it is *metaphysically* impossible. So the factualist needs the *metaphysical* possibility of something which isn't metaphysically possible.

But we don't need the metaphysical possibility of *anything.* What *we* do is merely to register the confirmation relations of some proposition, namely *Kennedy did not die in '63.* We don't need there to be a *world* that has the same laws as ours but where Kennedy didn't die. Such a world may well *not* be possible, in any but an epistemic sense. For if the laws in w are the same as ours, then any difference between our world and w that is sufficient for keeping Kennedy from dying is almost certainly going to be sufficient for a lot of other things that will make (v) be false — it might be sufficient for there not being a Kennedy or for Kennedy's dying of hemorrhagic fever in '63. All *we* need to make (v) true is the existence of some set of propositions $P_1 \ldots P_n$ — where *Kennedy did not die* is among them — such that those propositions give a high degree of credibility to the proposition *Johnson did not become President in '63.*

There is, I grant, more to it than that. (To facilitate discussion, let us suppose that P_n is the proposition *Kennedy did not die in '63.*) We cannot, of course, allow just *any* proposition to be included in $P_1 \ldots P_{n-1}$. It isn't clear what *all* of the selection criteria are, but it is clear what *some* of them must be. Trivially, $P_1 \ldots P_{n-1}$ cannot include the proposition *Johnson didn't become President in '63.* Less trivially, $P_1 \ldots P_{n-1}$ cannot include propositions *known* to be false; for example, P_3 cannot be the proposition *the black plague kills all American politicians (including*

Johnson) in '63. Also, $P_1 \ldots P_{n-1}$ cannot exclude propositions that we know to be true. Suppose we knew that Kennedy was dying of a rapidly progressing form of cancer which kills its victims in under three weeks. Obviously we couldn't exclude that from $P_1 \ldots P_{n-1}$. The rule seems to be that $P_1 \ldots P_{n-1}$ must include all and only known propositions.

But perhaps this rule is a bit simplistic. What a biologist considers counterfactually possible differs from what a physicist would consider counterfactually possible. If you ask a biologist whether a given zygote *could* have acquired a certain morphology in due course (say, that of a male), he may say "why of course : it is completely open which gender a zygote assumes — so of course it *could* have acquired that morphology". But if you ask a physicist the same question he may say : "given the initial positions and momenta of the atoms composing that zygote and composing the neighboring physical systems [and so on], it was inevitable that it would eventually assume that morphology." The biologist seems to be saying : nothing that falls within the scope of *biology* rules out the possibility that the zygote would become a female, not that nothing *tout court* rules it out. The physics is saying : something that falls within the scope of *physics* rules out the zygote's becoming a female.

Similarly, if you ask an economist and a historian what was counterfactually possible — whether Saddam *would* have done such and such *if* thus and such — they may give different answers ; for the historian will relativize his statement to propositions falling within the scope of his broadly historical concerns, whereas the economist will consider a slightly different data set.

But this relevance constraint might well reduce to the fact that the physicist *knows* certain things that the biologist does not and that the economist *knows* certain things that the historian does not (and perhaps vice versa). So it may be true that the one selection criterion is : $P_1 \ldots P_{n-1}$ must include all and only those propositions known (or reasonably believed) to be true by the person affirming the counterfactual (excluding, of course, the proposition whose negation is the antecedent of the counterfactual).

And it is clear that, given any one person, a great deal is epistemically possible from his perspective that is not *objectively* (metaphysically) possible. It was epistemically, but not metaphysically, possible that Hesperus should be distinct from Hesperus. Given only what *I* know, it is epistemically possible that the laws of physics remain exactly what they are *and* that there will not be an earthquake in California in the next year. But it might well be a *logical* impossibility — not a nomic or even a mere metaphysical impossibility — that the conjunction :

(*)$P_1 \ldots P_n$ and $C_1 \ldots C_k$ and *there is no earthquake in California in the next year*

should be true, where $P_1 \ldots P_n$ is the set of propositions giving the laws of physics and $C_1 \ldots C_k$ is the set of propositions describing the current condition of the California landmass. (*) may well be logically false because, given the relevant conditions and laws, it may follow deductively that there is an earthquake in California in the next year. In fact, this is the case if, in fact, there will be such an earthquake and the laws in question are deterministic.

For *our* analysis of counterfactuals to work, it is necessary only that *relative to what is known*, some untrue proposition, conjoined with propositions embodying the just mentioned knowledge, give high credibility to some other proposition. We don't need some set of initial conditions that didn't hold to *actually* hold, albeit in some other world, all the while making sure that *our* laws apply in that other world. All *those* desiderata cannot be jointly satisfied.

All *we* need is for some proposition (the consequent of the counterfactual) to be epistemically possible *relative to some false proposition* (the antecedent of the counterfactual) conjoined with *what we know* (minus our knowledge that the antecedent is false). And that is an extremely liberal constraint. For what is possible relative to what we know is extremely inclusive. It is possible, relative to what I know (minus my knowledge that Johnson *did* become President in '63) conjoined with the proposition *Kennedy didn't die in '63*, that Johnson did not become President in '63. In other words, if $\mathrm{p}_1 \ldots \mathrm{p}_m$ are the propositions embodying what I know, minus the proposition that Johnson became President in '63, then the conjunction of $\mathrm{p}_1 \ldots \mathrm{p}_m$ and *Kennedy didn't die in '63* gives high probability to the proposition *Johnson didn't become President in '63*. And that is what it is for the counter-factual *if Kennedy didn't die in '63, Johnson wouldn't have become President in '63* to be true.

If I were an extremely knowledgeable historian, it might *not* be the case that, relative to what I know, *Johnson didn't become President* was epistemically possible for me. But that is consistent with the nature of counterfactuals. The more a person knows, the fewer counterfactuals he can coherently entertain. If you ask somebody who knows nothing about American history what was *possible* for the country in 1813, he will say "just about anything". If you ask somebody who knew exactly how much money was in the treasury, exactly what the President at the time was thinking, exactly where the troops were stationed, and so on, that person will probably have a different answer.

This point about possibility obviously maps onto one about counterfactuals. Assuming maximal rationality on the part of the respondents (so that we don't have to worry about incompetence on their part)[8], the question "what *would* have happened to the U.S. if(...)?" gets very different answers from different people depending on what those people know. This doesn't mean that one of them is right and the other is wrong. It means that the question is inherently epistemic. One of the respondents, because he knows too much, cannot (without contradicting himself) fit the antecedent of the counterfactual into the list of propositions he knows to be true. The antecedent isn't epistemically possible for him (given what he knows). So he rightly balks at the question. But the antecedent is epistemically possible for the *other* respondent; so he doesn't balk.

VIII. To facilitate further discussion, I'd like to introduce a notational convention. For any event or state of affairs S, we will let $P(S)$ be the proposition affirming S's existence.

There are many apparent problems with our analysis of counterfactuals. I'd like to say why, in my view, they are not actual problems for it[9].

> According to your analysis, the meaning of "if A were the case, then B would be the case" is :
> $P(A)$ (together with certain other propositions) gives a high-degree of credibility to $P(B)$.
> It seems to me that your analysis is vulnerable to an exact analogue of the argument you directed against the factualist analysis. If A were the case, it would obviously be caused by some pre-existing state of affairs C_1. And C_1, in its turn, would be caused by some pre-existing state of affairs C_2. And so on. Of course, for each i, C_i would be a state of affairs that

[8] We must make this assumption. It isn't worth considering the case where stupidity affects what answers the person gives. I am assuming that counter-factuals are *objectively* true or false — they are not mere expressions of how we feel. (Some, amazingly, have argued that they are expressions of feeling. I think that *ultimately* — despite his numerous disclaimers — Blackburn has essentially this view. See [Blackburn 1984, chapter 6] and [Blackburn 1993]. I am arguing that *correct* counter-factuals are statements of the form : $P_1 \dots P_n$ gives such and such credibility to P_{n+1}. Obviously I want to illustrate this point by discussing people who assign the *right* probability to P_{n+1}, given $P_1 \dots P_n$, not the ones who do it the wrong way.

[9] All of the subsequent problems I will deal with — in other words, the points made by the objector — are points actually made to me by a very astute reviewer at *Philosophia Scientiae*.

> did not *actually* obtain. For reasons you described at great length, the conjunction $P(A)$ *and* $P(C_1) \ldots P(C_n)$ would either give little or no probability to $P(B)$ *or* that conjunction would probability propositions that were not even counterfactually true. So your analysis fails for the very same reasons that the factualist analysis fails.

This point overlooks the very thing that distinguishes my epistemic analysis from the factualist analysis. Imagine the following. You and I are homicide detectives. We are trying to solve the murder of Smith. We did not witness the murder, and neither did anyone else. But we have data that incriminates Brown. Brown's finger-prints are found on the knife lodged in Smith's rib-cage. Specimens of Smith's hair and blood were found in Brown's car and home. And so on. Let $P_1 \ldots P_n$ be the propositions that describe all of the data that we've gathered so far, along with the propositions that embody our knowledge of general regularities and causal laws. And let $P(M)$ be the proposition : *Brown killed Smith.*

By themselves, $P_1 \ldots P_n$ give a certain degree of probability to $P(M)$. That probability is less than 100%, but is still quite high. Given *only* $P_1 \ldots P_n$, it is still open whether Brown murdered Smith. For $P_1 \ldots P_n$don't *entail* $P(M)$. Perhaps Brown is being framed, and the evidence in question was planted. Perhaps, unknown to the world, Brown has an identical twin — who has the same finger-prints and genetic profile as Brown — and that twin committed the murder. That is all well and good. But there is no disputing that $P_1 \ldots P_n$, *by themselves*, give a certain probability to $P(M)$. We don't need *new* data to decide whether $P_1 \ldots P_n$ probability $P(M)$. They obviously do probabilify it.

Of course, there are many propositions P_{n+1} such that $P_1 \ldots P_{n+1}$ give extremely low probability to $P(M)$. P_{n+1} could be the proposition : *Jones, not Brown, killed Smith.* This would be a trivial example. A less trivial example would be the case where P_{n+1} is the proposition : *Brown has a sociopathic identical twin who has the same genetic markers as Brown and also has a key to Brown's car and house and has a habit of framing Brown.*

But the proposition $P_1 \ldots P_n$ *give 95% probability to* P(M) is simply distinct from the proposition $P_1 \ldots P_{n+1}$ *give 95% probability to* P(M). In fact, those two propositions *must* have different truth-values, unless P_{n+1} is some completely irrelevant datum.

Of course, to solve the homicide, we do need new data. But we need new data *not* to decide whether $P_1 \ldots P_n$probabilify $P(B)$. They do probabilify it ; we know that. We need new data to decide whether *in fact*

$P(M)$ is true. But it is beyond dispute that, by themselves, $P_1 \ldots P_n$ probabilify $P(M)$.

Indeed, new empirical data will not have any bearing on *how much* $P_1 \ldots P_n$ probabilify $P(M)$. New empirical data will bear on whether $P(M)$ is true or not. But the relationship between $P_1 \ldots P_n$ and $P(M)$ is a purely logical relationship. Granted $P_1 \ldots P_n$ do not *entail* $P(M)$. But the proposition : *if* $P_1 \ldots P_n$, *then* $P(M)$ *has a 95% likelihood of being true* is not itself an empirical proposition. $P_1 \ldots P_n$ is an empirical proposition. And so is $P(M)$. But the conditional proposition *if* $P_1 \ldots P_n$, *then* $P(M)$ *has a 95% likelihood of being true* is not empirical[10].

When one makes a probability-assertion, one is making a statement about the bearing of what one knows on some proposition not yet known to be true. One is *not* making a statement about the bearing of the *totality* of true propositions on some other proposition. Socrates either had two kidneys or he didn't. And yet the statement *it is 98% likely that Socrates had two kidneys* may well be true. (More exactly, it may be elliptical for a true statement.) That statement does not mean : *the conjunction of all true propositions gives 98% probability to the proposition that Socrates had two kidneys.* Probability statements *never* have that form. For the conjunction of all *true* propositions will either give 100% or 0% probability to any given proposition, depending on whether the given proposition is true or false. And, except to illustrate some purely theoretical point, we are not interested in such cases. Probability statements *do* have the form

(15) *if* P, *then the probability of* Q *being true is such and such.*

If we *add* propositions to the antecedent, we get an entirely different probability statement. If we *add* $P_2 \ldots P_{37}$ to the antecedent, then we end up with the statement

(16) *if* P *and* P_2 *and* $(\ldots)P_{37}$, *then the probability of* Q *is such and such.*

But (15) is a completely different proposition from (16).

My analysis of counter-factuals is this : *if* A *had been the case, then* B *would have been the case* means : The proposition

(17) K *and* $P(A)$ *gives a high probability to* $P(B)$,

where K is the conjunction of propositions we know to be true, minus the proposition that A *and* B *were not the case.* Of course, (17) is equivalent to :

[10]This crucial point about probability statements is made and decisively argued for in [Bonjour 1998].

(18) *if K and P(A) are the case, then it is highly probable that P(B) is true.*

If we *add* propositions to the antecedent of (18), then we simply end up with a *different* proposition from (18). The objector says that we must add such antecedents. But the truth is : we must *not* to do so. And this obligation *not* to do so lies at the heart of the concept of probability. So there is no regress of any kind, and no analogue of the argument we used against the factualist is generated.

IX. Here is another objection to our analysis :

> An alternative world is usually thought of as a set of propositions. Let's suppose this is right. In that case, I don't see how your view differs from that of the factualist. You say that a counter-factual registers some relation — presumably one of entailment or confirmation — holding among propositions. According to the factualist, a counter-factual says what is the case in some other world. If that other world just *is* a set of propositions, then he is really affirming the existence of some kind of relationship – presumably one of entailment or confirmation — among propositions. He is saying that, if $P_1 \ldots P_n$ are true, then so (probably) is some other proposition Q. (Among $P_1 \ldots P_n$ is the proposition affirming the antecedent of the counter-factual. The remainder are those propositions that describe our world *minus* the ones that are inconsistent with the proposition affirming the antecedent of the counter-factual.)
>
> Basically, if worlds just *are* sets of propositions, then the factualist approach collapses into your epistemic approach. And your approach therefore cannot possibly be superior to the factualist approach.

There are two problems with the objector's point. First, worlds are not sets of propositions. Second, while it is true that many factualists identify worlds with sets of propositions, the sets they have in mind — so-called "maximal consistent sets" — are absolutely useless for analyzing counter-factuals.

When I use the term "world", I mean what is ordinarily meant by that term : I mean the kind of thing that *this* world is. Our world is not a set of propositions. Our world is something which makes sets of propositions be true or false. Any world deserving of the name is something which makes propositions be true or false. A set of propositions does not make

propositions be true or false. The set of propositions <snow is white, grass is green, penguins fly> doesn't make any of those propositions be true or false. That set exists regardless of whether penguins fly, regardless of whether grass is orange.

Granted, there are may who maintain *both* that counter-factuals register facts about other "worlds" *and* who say that other "worlds" are just sets of propositions. Let us refer to such people as "weak factualists". (A "strong factualist" would be somebody, like David Lewis, who is a factualist and who holds that alternative worlds are concrete entities, not sets of propositions.)

There are many weak factualists. In fact, so far as I know, *every* one who writes about counter-factuals, apart from Lewis, is a weak factualist. But weak factualists never, absolutely never, hold that counter-factuals register relations of probabilification among propositions. In my view, that is what they *should* hold. But it isn't what they hold.

What do weak factualists hold ? It seems to me that what is going on is this. They start out by accepting Lewis' analysis. But they are not willing to go so far as to say that there really are alternative worlds, in any literal sense of "world". But they want to hold onto the Lewisian analysis, presumably because that analysis obviously does an excellent job of capturing our pre-theoretical understanding of counter-factuals. So, on the one hand, the weak-factualist isn't willing to completely give in and treat counter-factuals as inter-propositional relations. (That is why, prior to this paper, no factualist has ever said flat out that *if Kennedy had not been shot, Johnson would not have become President in '63* holds entirely in virtue of some confirmation or entailment relation holding among propositions.) But on the other hand, the weak-factualist isn't willing to go in the other direction : he isn't willing to say that *there really are* other worlds. *A fortiori* he isn't wiling to say that concrete occurrences are what make counter-factuals be true. So the weak-factualist denies that counter-factuals merely register relations of probabilification (or entailment) among propositions. *And* he denies that counter-factuals register facts relating to other worlds (in any genuine sense of "world"). So his position is completely incoherent.

If we are to understand counter-factuals, we must completely jettison the notion of alternative worlds. If by "world" we mean something at all like our world, then counter-factuals are not to be understood in terms of worlds. If by "world" we just mean "set of propositions", then we should just talk about sets of propositions, and not let a misleading bit of terminology create the illusion of explanation.

There is more to say. Apart from Lewis, everyone who uses the notion of other worlds identifies worlds with sets of propositions. But they don't ever say that a world is just *any* set of propositions. They always identify worlds with what are called *maximal consistent sets*. What is a maximal consistent set of propositions ? S is a maximal consistent set of propositions iff, for any atomic proposition P[11], either P or its negation is included in S.

Why are worlds identified with *maximal consistent* sets ? The reason is clear. Consider the following set of propositions. <Grass is green, penguins fly>. This set of propositions obviously doesn't determine a world ; too much is left out. Any number of *different* worlds could accommodate this set. A world is supposed to something which is determinate in every respect. The set <Grass is green, penguins fly> leaves too many things indeterminate. Suppose, for the sake of argument, that that set were a world — call it w. In that case, the following questions would have no answer. "Are there tigers in w ? Are there nine planets in the solar system in w ? Did Keats live to be 90 in w ?" Obviously w is much too thread-bare to count as a world : a world is something which gives answers to questions.

So if worlds are, with any plausibility, to be identified with sets of propositions, they must be sets of propositions that don't leave too many questions unanswered. To turn w into something that is even minimally deserving of the title "world", we have to flesh it out — we have keep adding propositions to it until w becomes capable of settling all the questions we might pose about it. This means that, for any question Q we might have about w, one of the propositions in w must be an answer to Q *or* the propositions that are in w must *entail* an answer to Q.

Now we can see why alternative worlds, in the sense of maximal consistent sets of propositions, are absolutely useless for analyzing counterfactuals. Consider the counter-factual

(19) *if Kennedy had not been assassinated, then he would have been President until '68.*

Here is what the weak-factualist says :

> (19) is true because in some world w, Kennedy was not assassinated and, in w, he remained President until '68. w is a set of propositions. w comprises the proposition *Kennedy was not assassinated*. With that qualification, w is as much like the set of propositions that describe *our* world as any

[11] An atomic proposition is one that is not a generalization of some kind and is not built out of other propositions.

> consistent set of propositions could be. w also comprises the proposition *Kennedy was President until '68*.

But, in effect, we've already seen the problem with this view. If w is to comprise the proposition *Kennedy was not assassinated*, then either w is will be replete with inconsistencies *or* w will describe a world so utterly different from our world that it almost certainly does not comprise the proposition *Kennedy was President until '68*.

Remember what we said about the strong-factualist position. If w is a world, in the literal sense, in which Kennedy is not assassinated, then w will differ from our world in so many ways that, in w, it is exceedingly unlikely that Kennedy is President until '68 : under those circumstances, it is unlikely that Kennedy even exists.

This fact about w corresponds to a fact about the set of propositions *describing* w. That set must include *Kennedy was not assassinated*. But if that set is to be consistent, then it must be dramatically different from the set of propositions which describes our world. But if the former set is *that* different from the latter, then it is, at best, minimally likely that it also comprises the proposition *Kennedy was President until '68*.

All the points that we earlier made about worlds are really points about propositions describing worlds. The position of the weak-factualist is very close to that of the strong-factualist. The only difference is that the weak-factualist talks about the propositions *describing* worlds, whereas the strong-factualist talks about worlds themselves. But all of the problems bedeviling the strong-factualist position are problems with the proposition *describing* worlds : there is no way to insert the antecedent of a counter-factual into that set without making the set prohibitively different from the set of propositions which describes our world. So the weak-factualist position is barely any better than the strong-factualist one.

The problem with the weak-factualist position is that maximal consistent sets comprise *too many* propositions. A maximal consistent set answers *too* many questions. Among those questions are questions relating to nomic relations among propositions. So if a consistent set is going to comprise a proposition affirming some proposition that is not true of our world — some proposition that could be the antecedent of a counter-factual — then one of two things holds. (a) That set will describe a world that bears no resemblance to ours. (b) That set is going to have to leave a lot of questions unanswered.

If (a), then the set is useless for running counter-factuals. If (b), then the set is not a maximally consistent set, but only some tiny subset of

such a set. So the weak-factualist ends up identifying worlds with tiny subsets of maximal consistent sets of propositions.

I grant that *if* he does this, then his position can work. But if he does that, then his position does not have anything to do with "worlds" at all, not even relative to the most permissive interpretation of that term. Everyone who talks about possible worlds is either talking about concrete worlds or about *maximal* sets of propositions. On either reading, the correct analysis of counter-factuals cannot possibly have anything to do with possible worlds.

X. These points enable us to deal with another objection :

> You say that counter-factuals register relations of some kind holding among propositions, not among worlds. These days propositions are generally held to be sets of worlds or, at any rate, functions involving worlds. If this conception is the right one, then I don't see the difference between your "epistemic" analysis and the traditional "factualist" one. Both analyses, apparently, turn counter-factuals into relations holding among propositions.

Again, we must distinguish definitions from analyses. It is, indeed, widely held that propositions are sets of worlds, or functions involving worlds. But this is not a definition of "proposition" ; it is a controversial *analysis* of what propositions are.

The uncontroversial *definition* of "proposition" is this : propositions are those things that are non-derivatively true or false. Sentences are true or false, but only because they bear propositions. ("Snow is white" is true entirely because it bears the proposition *snow is white.*) Beliefs are true or false, but only because they encode propositions. (My belief that snow is white is true entirely because it encodes the proposition *snow is white.*) Whenever we have something that is true or false, we either have a proposition *or* we have something which inherits its truth-value from a proposition.

When I use the term "proposition", all I mean is : something which is non-derivatively true or false. I mean the kinds of things that are believed and expressed by sentences.

To facilitate discussion, let "possible world semantics" (PWS) be our term for the view that propositions are sets of worlds or functions involving worlds.

The objector says that *if* PWS is correct, then my analysis is not (appreciably) different from the factualist one. I must disagree.

If someone advocates PWS, he cannot then say that worlds are themselves sets of propositions. For then his analysis of "proposition" is this : a proposition is a set of sets of propositions or is a function from sets of sets of propositions to truth-values. Obviously such an analysis is viciously circular.

So someone who advocates PWS is quite definitely committed to the view that worlds are *not* sets of propositions. He is therefore committed to the view that worlds are concrete objects.

As we have seen, if worlds are concrete objects, then the factualist position fails. So contrary to what the objector says, the factualist must hold that propositions are *not* sets of worlds or functions involving worlds.

But even if we leave aside the last point, the objector's point fails. For the sake of argument, suppose the objector is right about propositions : suppose that, indeed, propositions are sets of worlds or functions involving worlds. Within this framework, how are we to analyze the notion of entailment ? In other words, supposing that PWS is right, what is it for P to entail Q ?

The only answer that is possible within this framework is this : P entails Q iff every "world" in which P is true is a world in which Q is true[12]. In other words, Q is the case in every world in which P is the case.

But this answer is not the right one. *Water is wet* is true in every world in which H_2O *is wet* is true, and *vice versa*. But no entailment relation holds between the two propositions. A relation of necessitation holds, but not one of *entailment*.

Of course, an advocate of PWS will try to amend that doctrine, so that it it will recognize the distinction between entailment and necessity. I wish to show that no such amendment is possible. Let us start with some truisms.

For P to entail Q is simply for *not [P and not Q]* to be a logical or analytic truth. (This is a truism, not a controversial analysis.)

There are no worlds in which the following is true : x is water and is not H_2O. According to PWS, that is why *x is water* necessitates x is H_2O.

There are no worlds in which x is a triangle and is does not have three sides. According to PWS, that is why *x is a triangle* necessitates *x has three sides*.

[12] And this is precisely the answer given by Hintikka, Carnap, and countless others.

But *x is a triangle* not only necessitates *x has three sides*. The first *entails* the second. By contrast, *x is water* does not entail *x is* H_2O.

How does PWS explain this difference? PWS has only a couple of options. One of these is to say :

> (E) There are *epistemically* possible worlds where the proposition [*x is water and x is not* H_2O] is true. And because such worlds are *epistemically* possible, one is not guilty of incoherence in believing that proposition.

But (E) is doubly incoherent. First of all, there are *no* worlds where that proposition is true. *A fortiori* there are no epistemically possible worlds where it is true.

Granted, that proposition is true in some worlds, if by "world" we mean only "set of propositions. But as we saw earlier, this *cannot* be what an advocate of PWS means by world. For if this is what he means, then his analysis of *proposition* becomes : a proposition is a function from sets of sets of propositions to truth-values. To avoid vicious circularity, the advocate of PWS must have *concrete* worlds in mind. There are no *concrete* worlds where the proposition [*x is water and x is not* H_2O] holds. So, in effect, there are *no* worlds where it holds.

There is a second reason why (E) is incoherent. As we saw earlier, the notion of epistemic possibility is to be defined in terms of the notion of self-contradiction. So PWS cannot use that notion to analyze the notion of entailment or, therefore, of self-contradiction.

We've considered one way that PWS might try to analyze the notion of entailment. And we've seen that this way fails. What other options does PWS have? It could say this :

> There are no worlds where [*x is water and x is not* H_2O] is true. And there are no worlds where [*x is a triangle and x does not have three sides*] is true. Nonetheless, one isn't guilty of incoherence if one believes that there are worlds of the former kind. But one is guilty of incoherence if one believes in worlds of the second of kind. That is why *x is a triangle* entails *x has three sides*, whereas *x is water* does not entail *x has three sides*.

But if he says this, the proponent of PWS is giving up on the idea that entailment can be analyzed in terms of relations among worlds. For he is, in effect, saying that the relevant relation among worlds must *itself* be understood in terms of entailment.

Uncontroversially, P entails Q iff [P and not Q] is self-contradictory and iff not [P and not Q] is logically necessary. (This is an uncontroversial truism, not an analysis.) The objector says : P entails Q iff it is self-contradictory to believe in a world where [P and not Q] is true. This is tantamount to : [P and not Q] is self-contradictory iff it is self-contradictory to believe in a world where [P and not Q].

This last point is true enough. But it is no *analysis* of the notion of self-contradiction. It *uses* that notion ; it *presupposes* it. It is therefore no analysis of it.

PWS completely and utterly fails to elucidate the notion of entailment. Whenever we try to analyze entailment in terms of relations among worlds, we end up presupposing that very notion. P *entails* Q if, given P, Q is not *just* necessary, but *epistemically* necessary. To keep epistemic necessity from collapsing into metaphysical necessity, PWS must invoke the notion of *coherence*. But once PWS does that, then its analysis of entailment becomes viciously circular.

I submit that, because PWS fails to analyses to the notion of entailment, it must be taken to be a wholesale failure. The notion of entailment is intimately related to the notion of a proposition. In fact, a standard view is that a proposition P *is* the class of its logical consequences[13]. Even if that view is false, the relationship between the notion of entailment and propositionality is *almost* one of identity : as a matter of conceptual necessity, all and only propositions stand in entailment-relations.

Clearly the relation between these two concepts is *so* deep that, if a doctrine fails to analyze the concept of entailment, then it fails to analyze the concept of a proposition. PWS fails to analyze the former. We must conclude that it fails to analyze the latter : it doesn't give a correct of what propositions are. Propositions are *not* sets of worlds or functions involving worlds. The notion of a proposition is *not* to be analyzed in terms of the notion of a world. This fact blocks attempts to view my epistemic analysis as a version of the factualist analysis.

[13][Lewis C. I. 1946], [Carnap 1947], and countless others have held this. I myself hold a version of this view. It has to be refined a bit to avoid some set-theoretical paradoxes. But the basic idea is unquestionably correct. If P entails Q then there is obviously an *internal* relationship between P and Q : a relationship that couldn't possibly hold between two *distinct* things. See [Ramsey 1990/1927, 42]. So some kind of containment relationship must hold between them. In my book *Conceptual Atomism and the Computational Theory of Mind* (unpublished) I defend, at great length, this very conception of propositionality. And I show why, contrary to first appearances, that conception is consistent with the fact that entailments are often informative.

XI. I would like to bring up one consideration that favors our analysis. The factualist analysis of counterfactuals is characterized by a certain circularity. I will argue that the circularity in question is vicious.

Not all true universal statements express laws.

Consider the statement : *Every coin in my pocket is a quarter.* That is a true universal statement. For it is equivalent to :

(20) for any x in the universe, if x is a coin in my pocket, then x is a quarter[14].

But (20) does not express a law. Let us contrast (20) with a proposition which *does express* a law (albeit a derivative one) :

(21) Pinewood floats on water.

What is the difference between (20) and (21) ? The generally accepted answer is this : (21) supports counter-factuals ; (20) does not. If I *were* to put a nickel or a penny in my pocket, it wouldn't turn into a quarter[15]. So the counter-factual claim

(22) For any x, if x *were* a coin in my pocket, it would be a quarter

is false. And that is why (20) does not express a *law.* By contrast, the counter-factual claim :

(23) For any x, if x were a piece of pine-wood placed on water, then x would float

is *true.* And that is why it expresses a law. So what is a law ? Here is a partial answer :

(L) For something to be a law of nature, it is at least *necessary* that it support counter-factuals[16].

Here is the problem for the factualist. The factualist says that counter-factuals make statements about other worlds that are like this world, except in some very limited respect. When we ask the factualist how are those worlds like our world ?", he says (among other things) : "Those other worlds are governed by the same *laws* as ours." So the counter-factualist says :

[14]This example has been used many times. Goodman [Goodman 1954] and Hempel [Hempel 1966] use it to illustrate this very point.

[15]My choice of examples here was influenced by private discussions with Ernan McMullen.

[16]Ayer in [Ayer 1972, 132] writes : "We cannot, indeed, define the concept of natural law without making use of conditionals — for one thing, we want to allow for there being laws of nature which have not yet been discovered, and this requires us to conceive of there being true generalizations which we should treat as generalizations of law if we cam to believe them (. . .)" Also, see [Mackie 1973, 65] and [Hempel 1966].

(CF) A counter-factual makes a statement that concerns (or is made true by) states of affairs in worlds that are governed by the same laws that govern our world.

Practically *everybody*, factualists included[17], believes (L). So when we ask a factualist what a law of nature is, he says "it is something that supports counter-factuals". And when we ask him what a counter-factual is, he says "something that is made true by worlds that have the same laws as our world". This, it seems to me, is plainly circular.

Granted, not all circularity is vicious. In contemporary philosophy of science, it is a commonplace[18] that scientific theories are characterized by a *certain* circularity. A theory is (inter alia) a set of inter-connected, mutually supporting propositions. Of course, a theory has to have *some* roots in data lying outside that theory. But *within* a theory we may find many cases of circularity. (Propositions about protons will explain propositions about electrons, and *vice versa*.) The theory *as a whole* cannot be circular : it has to be anchored in, and answerable to, extra-theoretical facts. But *within* the theory, there are many cases of propositions giving explanatory support to *one another*. So, once again, we see that not all circularity is vicious[19].

Now we can see why the factualist is guilty of vicious circularity. Scientific theories are given by epistemically contingent propositions ; so they gain or lose support with new empirical information. What (L) and (CF) say has nothing do to with anything contingent. They are *conceptual* explanations — attempts at conceptual analysis. (CF) doesn't say : "Given the conditions and causal laws that *in fact* prevail, a counter-factual is such and such." (CF) isn't the kind of proposition that gains or loses support with new empirical data. Neither is (L). They are both meant to be non-contingent conceptual analyses. They may be right or wrong. But they are *not* propositions that are confirmed or disconfirmed by empirical data. So we cannot explain away the circularity in question by saying that it is the innocuous kind (supposedly) characteristic of scientific theories.

The reason empirical theories aren't (always) viciously circular is that such theories are given by epistemically contingent propositions. Therefore such explanations do not give conceptual analyses : they do not give

[17]The only philosopher I know of who denies (L) is Bas van Fraassen. But van Fraassen's conception of natural law and of counter-factuals does not involve the notion of possible worlds at all.

[18]Not one I agree with.

[19]In his paper "Theories", Frank Ramsey (supposedly) showed why the kind of circularity (supposedly) characteristic of empirical theories isn't vicious.

the *content* of propositions. They affirm the existence of (epistemically) contingent relations among propositions.

But conceptual explanation is a different story. Where that kind of explanation is a concerned, circularity means a loss of *content*. The proponent of (L) is trying to give a conceptual analysis of the concept of a law of nature : he is trying to give the *content* of that concept. The proponent of (CF) is trying to give a conceptual explanation of the concept of a counter-factual : he is trying to give the *content* of that concept. Somebody who puts forth both (CF) and (L) as conceptual explanations is incapable of even giving *content* to either proposition. (CF) has content only in so far as the concept of (L) has some independent content. And (L) has content only if the concept of a counter-factual has independent content. If each has content only on the condition that the other does, then neither has content. So it is not possible that both (CF) and (L) should constitute conceptual explanations. But that is exactly what they are supposed to be.

These points enable us to deal with another attempt to show that the circularity in the factualist position is innocent. We believe *both*

(a) bachelors are unmarried males

and

(b) unmarried males are bachelors.

But we are not, on that account, guilty of *vicious* circularity. There is an element of circularity in all tautologous beliefs. But one is not guilty of vicious circularity in virtue of having such beliefs. Perhaps, it might be thought, it is equally innocent to hold both (CF) *and* (L).

This is not so. (CF) is meant to be an *analysis* of the concept of a counter-factual : it is meant to give the *content* of that concept. (L) is meant to be an *analysis* of the concept of law : it is meant to give the *content* of that concept. So, for reasons discussed a moment ago, holding *both* (L) and (CF) strips them both of content. That is why the position is *viciously* circular.

If I put forth (b) as an *analysis* of (a), and I also put forth (a) as an *analysis* of (b), then I *would* be guilty of vicious circularity. But that is not what I do. I believe both (a) and (b). But my analyses of *male* and *unmarried* don't involve the concept *bachelor*. They involve *different* concepts — concepts relating to certain legal institutions, to possession of certain reproductive organs, and so on. That is why viciously circularity is avoided.

XII. On our view, there is no circularity at all. For counterfactuals are *just* probability statements, and probability is a notion that can be

defined totally independently of the notion of a law.

The truth of a probability statement supervenes entirely on what is known to be the case — not on what *would* be the case or what *must* (in a nomic sense) be the case. Suppose you say 'the probability that next ball drawn from the urn is .9'. The truth of that statement derives entirely from what *has* happened and, indeed, from what has been *known* to happen. If it is known that 90% of the balls thus far drawn from the vase have been white, then (ceteris paribus), that probability statement is accurate.

Of course, in many cases, knowledge of laws of nature must be taken into account : the statement "the probability that that bar of steel conducts electricity better than that piece of rubber is 99.9999%" is based on (inter alia) knowledge of laws of nature. But the basis for one's belief in those *laws of nature* ultimately reduces to one's knowledge of what *has* happened. So *ultimately* probability statements are warranted by what *has* happened. Suppose Bob is in world w and Bob* is in world w^* and the evidence available to Bob concerning what *has happened* (in w) is identical with that available to Bob* concerning what *has* happened (in w^*). In that case, Bob and Bob* cannot diverge in their probability-judgments unless one of them is being irrational (or, in any case, unless one of them is doing a worse job than the other in analyzing the data)[20]. So the concept of probability is definable entirely independently of the concept of law and the concept of what *would* have

[20]The obvious response is this :

Suppose that Bob and Bob* have the exact same beliefs concerning what *has happened*, but Bob knows about laws that Bob* doesn't know about. For example, Bob knows that metal expands when heated, Bob* doesn't know this, but their respective knowledge of what has *happened* — of what states of affairs have obtained — is the same. In that case, Bob and Bob* can *rationally* assign different probabilities to, say, the possibility of some metal's expanding upon being heated.

Remember that whatever probability can reasonably be attached to a putative law of nature, the *sole* basis for that evaluation must lie in what has happened : *ultimately* all probability-assessments flow from what *has* happened and, indeed, from what is known. Of course, Bob may believe in some law L *not* because he has personally witnessed states of affairs that confirm it, but because some scientist told him about it or because he read about it in a book written by an authority. But in this case, Bob's 'knowledge' is obviously parasitic on other people's knowledge of singular states of affairs. *In the end*, what gives probability to a position is the totality of known *singular* states of affairs. So the way to respond to the objector is to say : if we leave aside parasitic knowledge (like Bob's knowledge of the laws of physics) or, alternatively, if we take into account the singular knowledge had by the communities to which Bob and Bob* belong, and from which their knowledge of laws is derivative, *then* there cannot be a divergence in the probability-assessments of these two people *unless* one of them is being irrational or, in any case, unless one of them is evaluating the data less competently than the other.

happened (the concept of counterfactuality). So our analysis is not guilty of circularity.

While being non-circular, our analysis accommodates the fact that statements of law do, while accidental generalizations do not, support counter-factuals. *Pinewood floats in water* states a law and thus supports counter-factuals. According to our analysis, the statement : *if this piece of pinewood were put in water, it would float* means that the proposition : *this piece of pinewood is in water*, plus some other set of propositions $P_1 \ldots P_n$ (all of them true), gives very high probability to the proposition *this piece of pinewood is floating*. Among $P_1 \ldots P_n$ are propositions like : *all past pieces of pinewood have floated in water* and *all objects having a certain composition float in water*. The result is that *this piece of pinewood is floating* is given vanishingly close to 100% probability. So our analysis correctly predicts that *pinewood floats in water* supports counter-factuals and therefore is not an accidental generalization.

Now consider the universal statement : *every person in this room is a philosopher interested in non-Lewisian analyses of counter-factuals.* This is a true universal statement. But it doesn't state a law, for it doesn't support counter-factuals. *If Bjorn Borg*[21] *were a person in this room, he would be a philosopher interested in non-Lewisian analyses of counter-factuals* is a false counter-factual. According to our analysis, its meaning is : the proposition *Bjorn Borg is a person in this room*, conjoined with certain other propositions $P_1 \ldots Pn$, all of them true, gives extremely high probability to *Bjorn Borg is a philosopher interested in non-Lewisian analyses of counter-factuals.* But they do *not* give high probability to that assertion. In fact, they give it (vanishingly close to) 0% probability. (For among $P_1 \ldots Pn$ are propositions like *Bjorn Borg hasn't the slightest interest in counter-factuals.*) So our analysis correctly predicts that *every person in this room is a philosopher* does not sustain counter-factuals and therefore does not state a law.

I must consider a very clever objection to my analysis (due to an anonymous reviewer) :

> You've said over and over again that the concept of law is to be analyzed in terms of the concept of a counter-factual. But if it is correct, then your analysis of counter-factuals is deeply absurd.
>
> According to that analysis, counter-factuals are just probability- or entailment-statements. The counter-factual
>
> (*)*If S had been the case, then S* would have been the case*

[21] The famous tennis player.

> (**) means : $P(S)$, *conjoined with* $P_1 \ldots P_n$, *highly confirms* $P(S^*)$,
>
> where $P_1 \ldots Pn$ are the propositions one knows to be true, minus the proposition that S was not the case.
>
> So you analyze laws in terms of counter-factuals. But you analyze counter-factuals as purely logical relations holding among the propositions *we know* to be true. So, by implication, you are advocating a kind of idealism or relativism about laws. On your analysis, laws are facts about our *beliefs* about the world, not about the world itself. But obviously laws are facts about the world itself, not about our beliefs.

I agree that laws (or, more precisely, the statements affirming them) necessarily support counter-factuals. I also agree that *if* counter-factuals make statements about worlds — about concrete scenarios — *then* those worlds must be governed by the same laws that govern our world. But I *don't* think that counter-factuals make statements about worlds.

Consider the counter-factual :

(24) *If Kennedy had not been assassinated, then Kennedy would have been President until '68.*

It cannot be denied that *if* (24) says something about what happens in some other world w, *then* w must be governed by the same laws that govern our world. If the laws in w are sufficiently different, or even minimally different, then Khrushchev might be President in w between '63 and '68.

But I myself do *not* think that counter-factuals are to be analyzed in terms of the occurrences in other worlds or scenarios. On my view, *no* alternative worlds are required. Other worlds, and the laws that govern them, drop out as irrelevant. On my analysis, (24) is simply a probability statement whose antecedent consists of the propositions one knows to be true, minus the proposition *Kennedy was assassinated*, and whose consequent is another proposition. Further, as we saw earlier, the antecedent ultimately doesn't include any statements of law.

So I *don't* hold that counter-factuals are to be analyzed in terms of the concept of law. We thus break the circle which leads from my view to idealism about laws.

XIII. *Why Humean laws support counter-factuals*

I would like to end this paper by arguing against a point whose correctness has been taken for granted.

Hume's theory of causation has been heavily criticized. One of the more striking criticisms is this :

> Humean laws are not ties *between* states of affairs. They are just regularities. Therefore Humean laws don't support counter-factuals.
>
> What is the case doesn't tell us anything about what would be the case *unless* what is the case includes forces or ties that guarantee certain outcomes, given certain initial conditions. If Hume is right, there are no such ties. What is the case is confined to the events that occur : there is nothing beneath them that guarantee anything. So Humean laws don't support counter-factuals (See [Mackie 1973, 65]).

I agree that *if* counter-factuals make statements about other worlds, *then* Humean laws do not support counter-factuals. If laws are Humean, there is nothing *underneath* events : there are just the events. And the occurrences in other worlds, if there are such things, are at least as likely to falsify our counter-factuals as they are to validate them.

But, of course, I don't think that counter-factuals make statements about other worlds. A counter-factual says simply that one statement probabilifies another. So

(25) if the vase had been dropped from a fifty story building, and nothing blocked or retarded its fall, then it would have shattered

means :

(26) the proposition *the vase was dropped from a fifty story building and nothing retarded its fall*, conjoined with $P_1 \ldots P_n$, probabilifies the proposition *the vase shattered*,

where $P_1 \ldots P_n$ are the propositions one knows to be true, minus the proposition *the vase was not dropped from a fifty story building*. There is no doubt that (26) is a correct probability-statement. And it is a correct probability statement *regardless* of whether Hume is right about laws or not. So Hume's analysis of law is consistent with the fact that (25) is a true counter-factual[22].

[22]I do think that the Humean analysis is wrong. But, to my knowledge, my reasons for thinking this have yet to be mentioned in the literature. Hume says that laws are regularities involving certain kinds of states of affairs. But this presupposes that the notion of a state of affairs is a non-causal one. For obviously if states of affairs are themselves causal processes, then Hume's analysis becomes viciously circular. Causation is a regular concomitance of states of affairs. And states of affairs are causal processes.

So clearly Hume is supposing that states of affairs, and objects, are not causal processes. But this supposition is wrong, even incoherent. It seems to me part of

References

Ayer, Alfred Jules

1972 *Probability and Evidence*, New York : Columbia University Press.

Blackburn, Simon

1993 Morals and Modals, in *Essays in Quasi-Realism*, Oxford : Oxford University Press.

1984 *Spreading the Word*, Oxford : Clarendon Press.

Blanshard, Bland

1961 The Case for Determinism, in *Determinism and Freedom in the Age of Modern Science*, Sydney Hook (ed.), New York : Collier Books.

Bonjour, Laurence

1998 *In Defense of Pure Reason*, Cambridge : Cambridge University Press.

Carnap, Rudolph

1947 *Meaning and Necessity*, Chicago : University of Chicago Press.

Fraassen, Bas van

1980 *The Scientific Image*, Oxford : Clarendon Press.

Goodman, Nelson

1954 *Fact, Fiction, and Forecast*, Cambridge : Harvard University Press.

Hempel, Carl G.

---

our concept of a state of affairs (leaving aside completely instantaneous ones) that such things are causal sequences. So Hume's analysis is circular. This connects with another problem with Hume's analysis. Hume takes it for granted that temporal order is something that can be understood independently of the notion of causation. x causes y only if events relevantly like x are followed, in time, by events relevantly like y. But if the notion of temporal succession is a causal notion, then this analysis is viciously circular. It is, at the very least, an open question whether temporal succession is a causal notion. Given what we know today, it seems likely that it is. A standard definition of temporal priority is this : x precedes y iff some possible causal sequence (e.g., a beam of light) connects x to y. It seems to me (though I am speculating) that if there is a possible causal connection between x and y, then there is an actual one. Given that forces just are deformations of space-time, and given the fact that no two events are totally causally isolated from each other, it seems to follow that any two non-simultaneous states of affairs are connected by some kind of causal process. So it seems to me that x precedes y iff there is some causal process starting with x that ends with y. If this is right, then temporal priority is itself a causal notion, and Hume's analysis fails.

1966 *Philosophy of Natural Science*, Englewood Cliffs : Prentice-Hall.

Hintikka, Jaakko

1962 *Knowledge and Belief*, Ithaca : Cornell University Press.

Keynes, John Maynard

1921 *A Treatise on Probability*, London : Macmillan and Co.

Lewis, Clarence Irving

1929 *Mind and the World Order*, New York : Dover. 1947 *An Analysis of Knowledge and Valuation*, Cambridge : Harvard University Press.

Lewis, David

1973 *Counterfactuals*, Oxford : Basil Blackwell.

Mackie, J. L.

1973 *Truth, Probability, and Paradox*, Oxford : Clarendon Press.

Quine, Willard van Orman

1982 *Methods of Logic*, Fourth Edition, Cambridge : Harvard University Press.

Ramsey, Frank

1990/1927 Facts and Propositions, in *Philosophical Papers*, D. H. Mellor (ed.), Cambridge : Cambridge University Press.